Lead Scoring Case Study

Anitha Ammati

Apoorva Srinivas

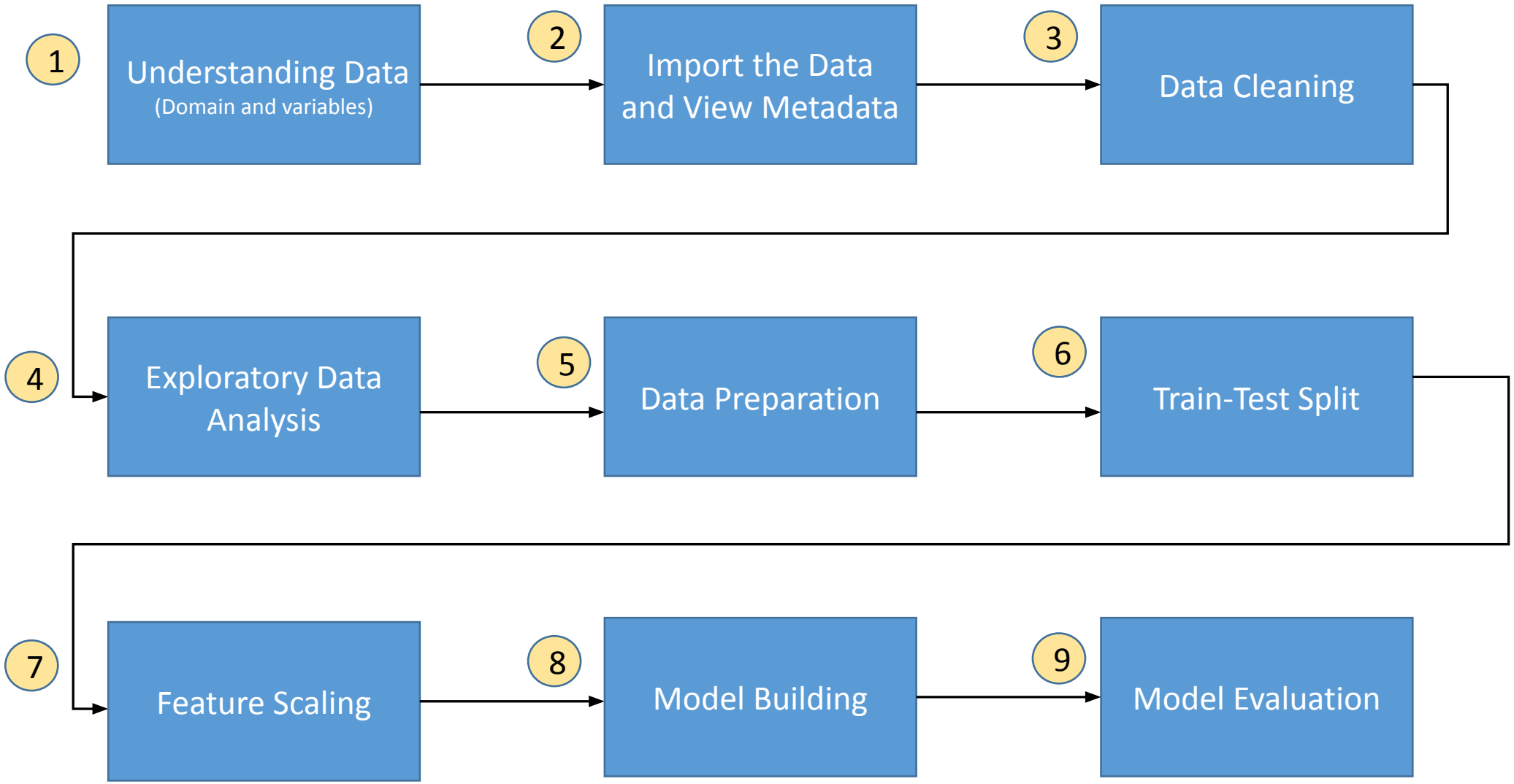
July 2021

Problem Statement

- An Education Company sells online courses to industry professionals.
- Many people visit the company's website through various marketing channels.
- The company has generated data from the people who have filled up a form on visiting their website. This data is used by the sales team for converting the leads into customers.
- The lead conversion rate of the company is 30%, which is quite small.
- Therefore, to increase efficiency, the company wants to generate a list of potential 'Hot' leads who are most likely to get converted to customers.

Typical Lead Conversion Process

Our Approach



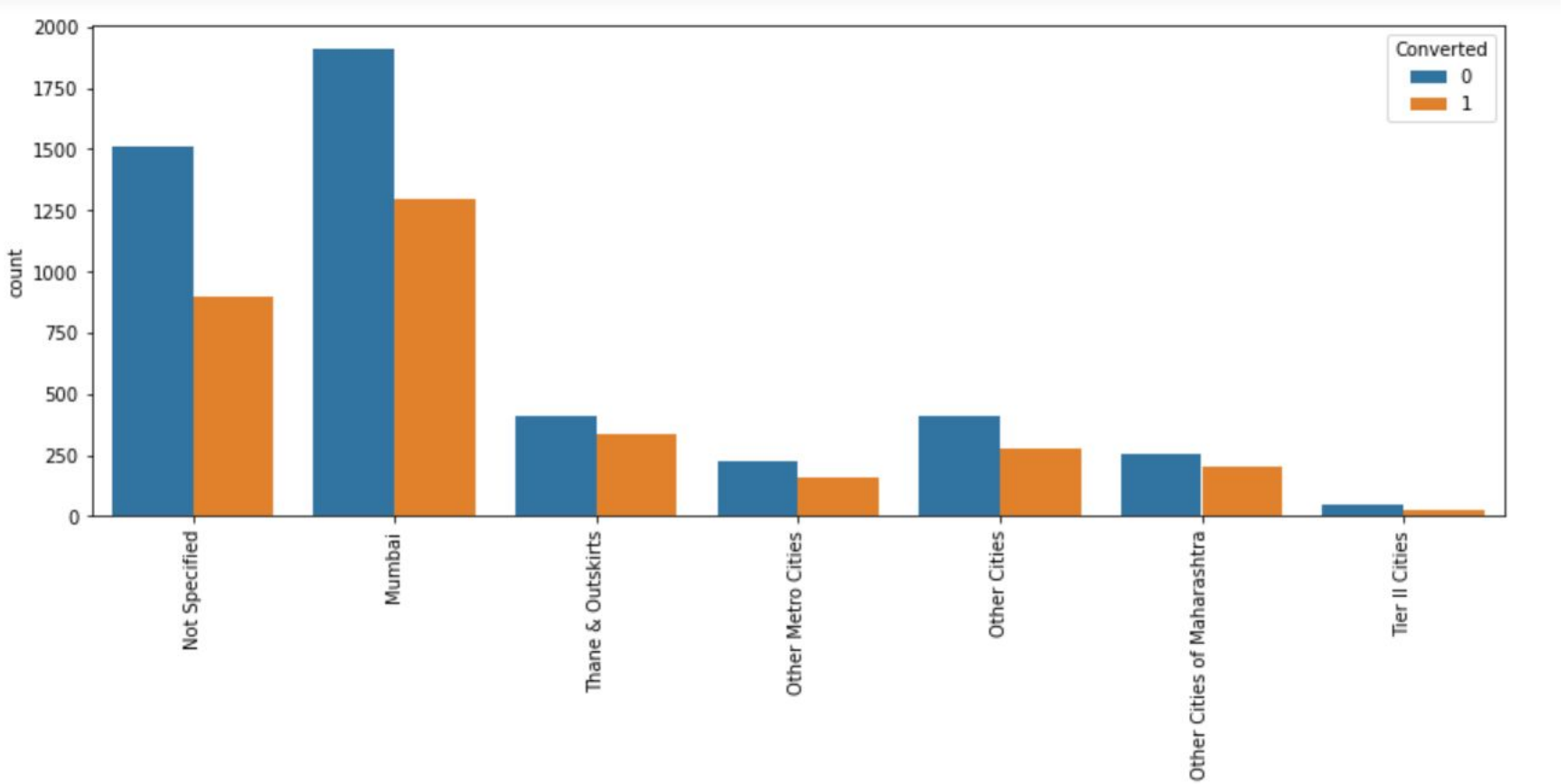
Details of Our Approach

The steps we followed in our analysis approach are as follows:

1. **Understanding Data:** The data dictionary helped us understand the variables and domain of the data.
2. **Import the Data and View Metadata:** In a Jupyter notebook, we imported the data as a Pandas Data Frame and viewed the metadata for basic analysis.
3. **Data Cleaning:** Performed missing value imputation and deletion of a few columns which were not useful for our analysis.
4. **Exploratory Data Analysis:** Univariate and Bivariate Analysis of the variables. This step provided more insight into the data.

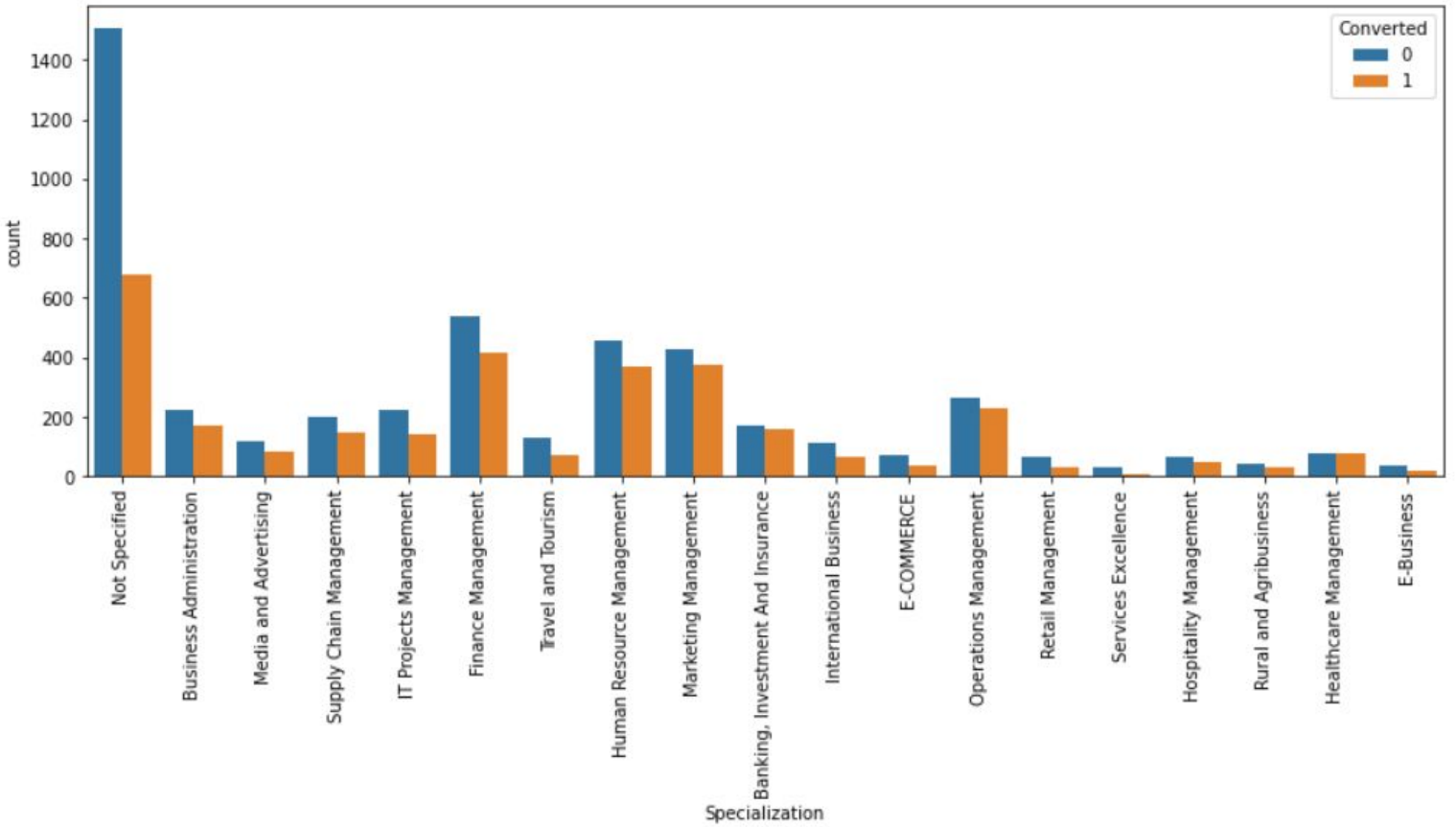
- **Data Preparation:** We handled outliers and class imbalance in the variables.
- **Train Test Split:** The given data was split into Training set and Testing set of the ratio 70:30.
- **Feature Scaling:** The features were standardized using the StandardScaler for consistency and to help the ML algorithm to converge faster.
- **Model Building:** The features for the algorithm were selected using RFE technique. The features which further proved insignificant were removed and the model was built again. The final model consists of 13 features. Prediction was made on the test set.
- **Model Evaluation:** The model was evaluated on various metrics such as Accuracy, Sensitivity, Specificity. The ROC curve showed accurate test result.
- Finally, a lead score in the range 0-100 was calculated for each row in the test set.

Observations



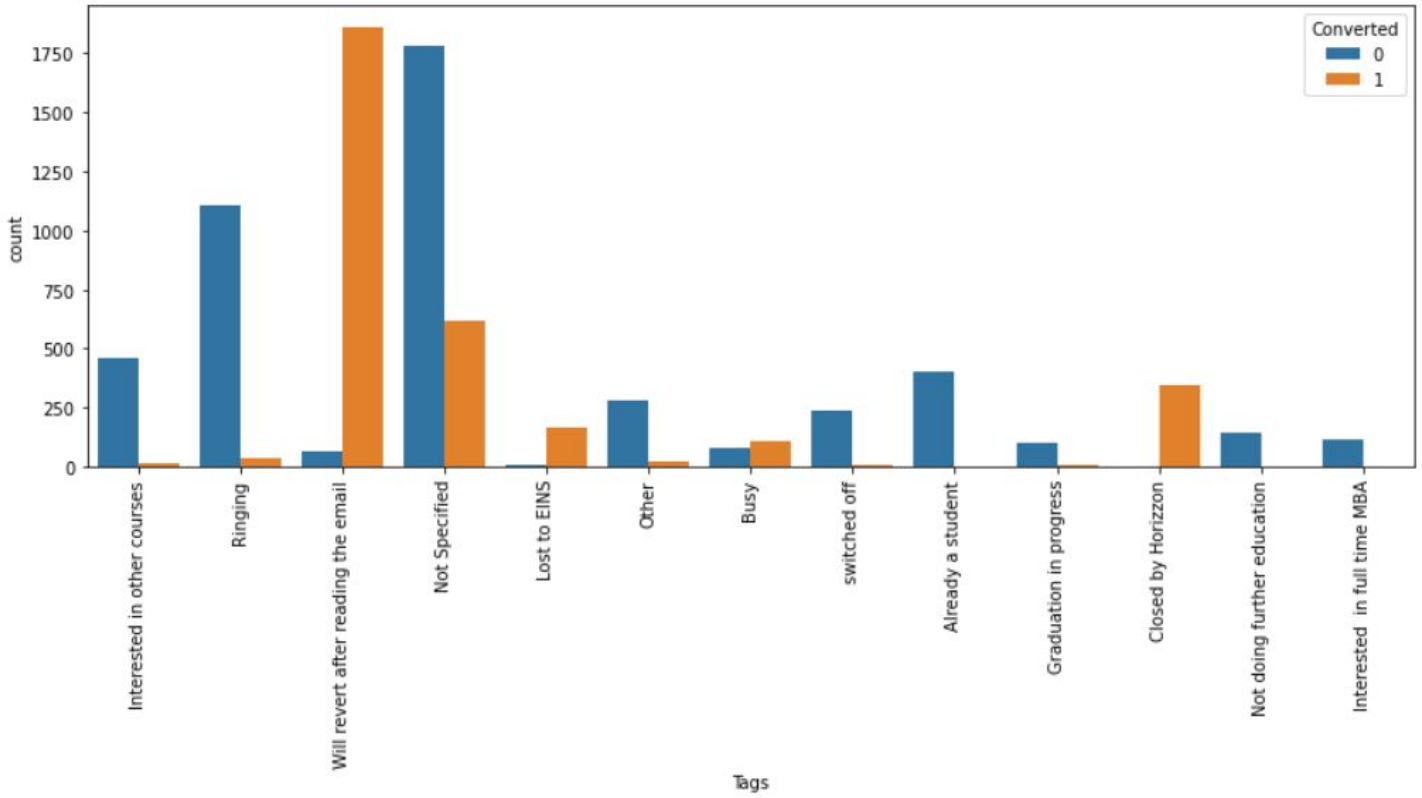
- Most of the leads live in the city of Mumbai.

Observations



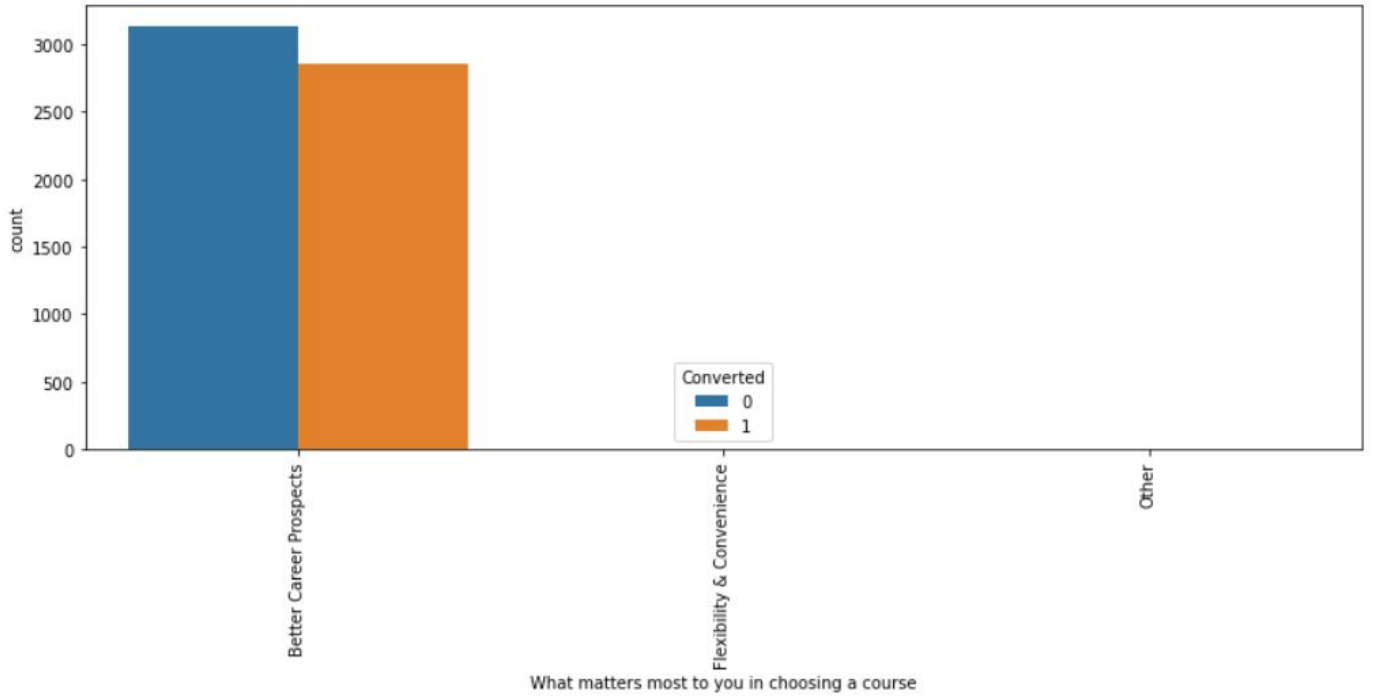
- Most of the leads have not chosen any specialization while filling the form.

Observations



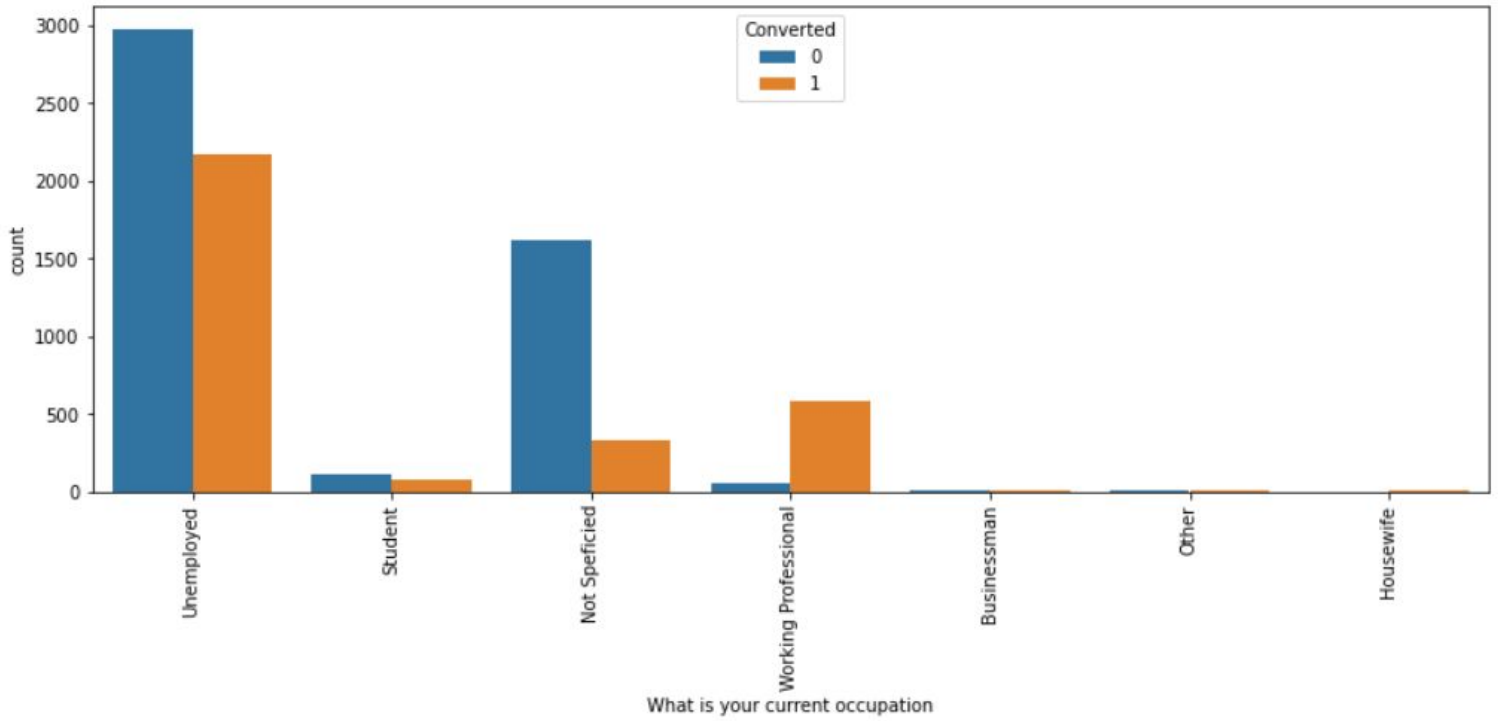
- The conversion rate is the highest for those leads who have indicated that they will revert after reading the email.

Observations



- Majority of the leads choose to take up an online course for better career prospects.

Observations



- Most of the leads who have expressed interest in the course are currently unemployed.

Summary

- A lead having a lead score of 30 and above have a higher chance of getting converted to a customer.
- Most of the converted leads are unemployed people looking for better career prospects.
- Initially, the converted leads are unclear about the specialization to choose. But, after detailed communication, they are most likely to enroll for the course.
- The tags given by the internal team for the leads play a significant role for lead scoring.
- The leads who have indicated that they will revert after reading the email are most likely to get converted.

Recommendations

Do's:

- The company must target leads with a lead score of 30 and above with aggressive marketing for conversion.
- The company must concentrate on leads who are tagged with the following: '*Lost to EINS*' and *'Will revert after reading the email'*.
- The leads who visit the company website are most likely to get converted. Therefore, the company must be proactive in updating the website and keeping it up-to-date.
- The Marketing Team must target advertisements on job portals where unemployed people frequently visit for job search.
- Since the courses are a success in Mumbai, the company could try to expand their reach to other cities in India.

Recommendations

Don’ts:

- The company should identify the leads whose phones are switched off or always ringing.
- The company must also refine the list of leads by excluding those leads where the emails have bounced.

The above categories of leads have a very low conversion rate and identifying them in the beginning helps in saving the company’s time and resources.

Recommendations

Subjective scenarios –

- If the company has the resources and time to spare, they could target leads with a lead score between 20 and 30. This might help in some leads getting converted to customers.
- In contrast, if the sales team has met their target and do not want to waste their time on unwanted calls, they must target leads whose lead score is 90 and above. This will ensure >90% of conversion rate.

Thank You